Estimation of state-space models for affective dynamics using Markov chain Monte Carlo methods

Tom Lodewyckx, Francis Tuerlinckx, Peter Kuppens
KULeuven

Nicholas Allen
University of Melbourne

Lisa Sheeber
Oregon Research Institute

Outline

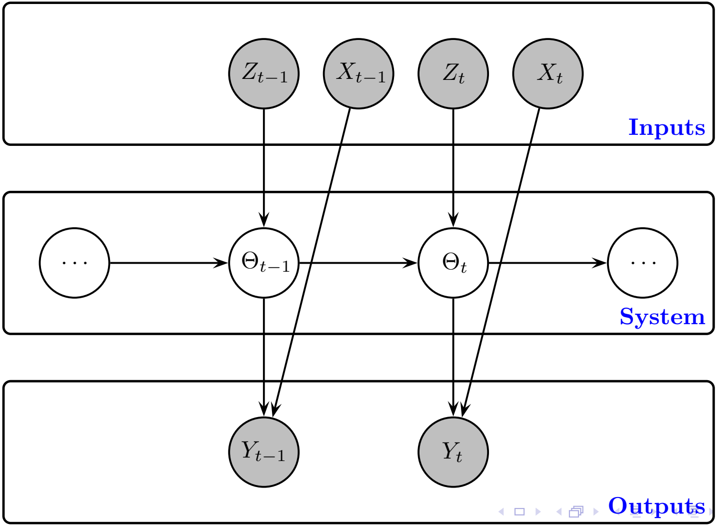
Outline

Introduction

Linear dynamical systems theory

- *Inputs:* observed noise variables / control variables
- *System:* unobserved variables, “states”, that have certain dynamical properties
- *Outputs:* observed variables of interest

Introduction



Introduction

Goal

- *Statistical framework:* Discrete-time state-space models
- *Conceptual framework:* Psychophysiological processes

  $\Rightarrow$ Combine these frameworks

Outline

Oregon adolescent interaction data

Experimental context

- 72 depressed, 69 normal adolescents
- Adolescent, father and mother in laboratory
- Perform interaction tasks
- Physiological measures adolescent
- Interactions videotaped, behavior coding system

Oregon adolescent interaction data

Experimental context

- 72 depressed, 69 normal adolescents
- Adolescent, father and mother in laboratory
- Perform interaction tasks
- Physiological measures adolescent
- Interactions videotaped, behavior coding system

Focus

- Problem solving task, unpleasant (pocket money)
- 18 minutes, measurements each second ($n = 1080$)

Oregon adolescent interaction data

Physiological measures (continuous)

- Heart rate (HR)
- Respiratory sinus arythmia (RSA)
- Skin conductance level (SCL)

Oregon adolescent interaction data

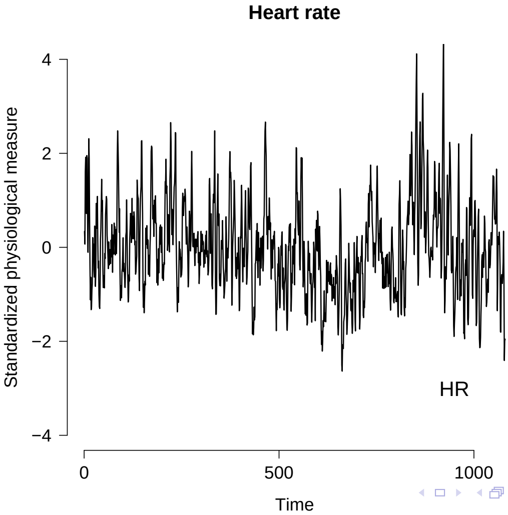
Physiological measures (continuous)

- Heart rate (HR)
- Respiratory sinus arythmia (RSA)
- Skin conductance level (SCL)

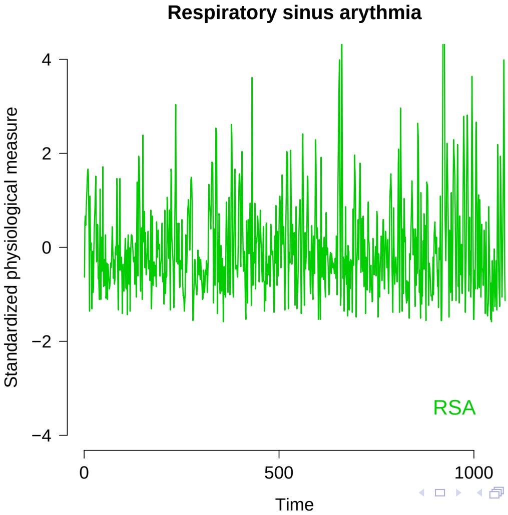
Behavioral measures (discrete)

- Anger adolescent (AdAnger)
- Anger father (FaAnger)
- Anger mother (MoAnger)

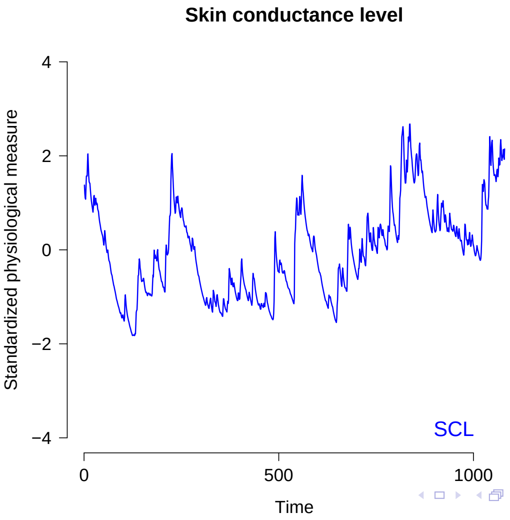
Oregon adolescent interaction data



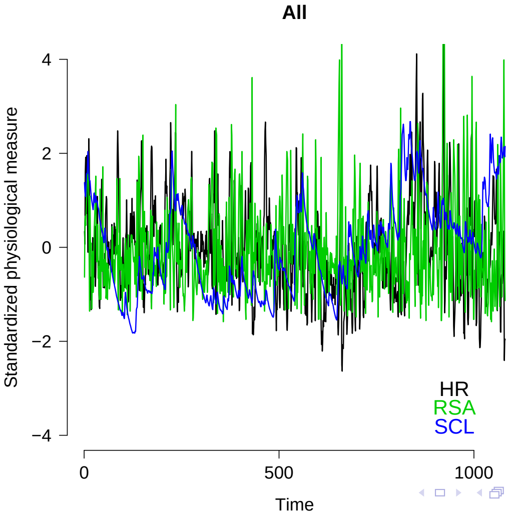
Oregon adolescent interaction data



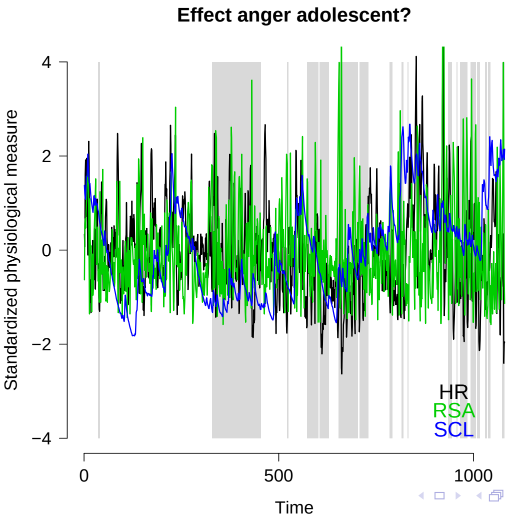
Oregon adolescent interaction data



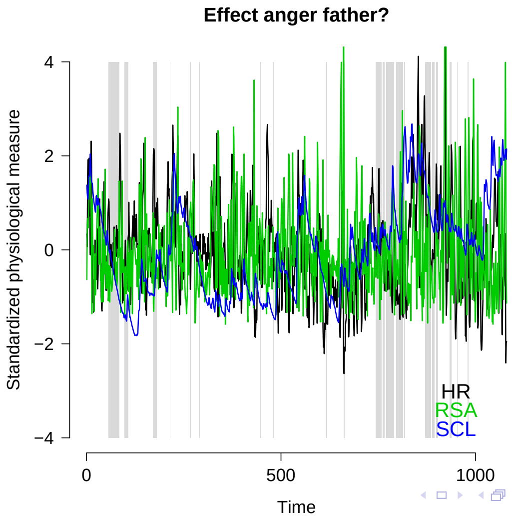
Oregon adolescent interaction data



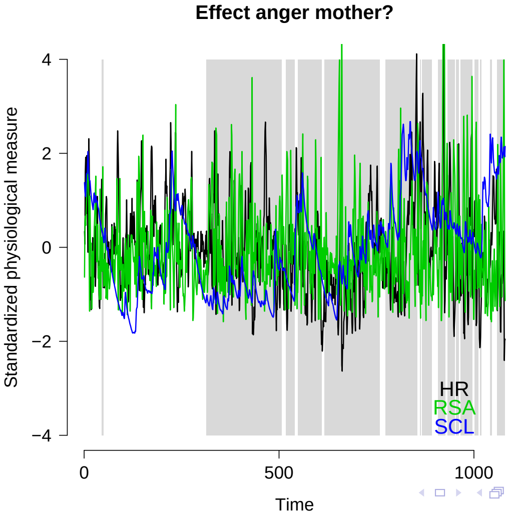
Oregon adolescent interaction data



Oregon adolescent interaction data



Oregon adolescent interaction data



Outline

Multivariate linear Gaussian state-space model

Application fields

- Engineering
- Econometrics
- ...

Multivariate linear Gaussian state-space model

Application fields

- Engineering
- Econometrics
- ...

Goals

- *State estimation:* True position of a space craft?
- *Parameter estimation:* Dynamical properties emotion system?

Multivariate linear Gaussian state-space model

Transition equation

$$Y_t \sim \text{Gaussian}\left(\Psi\Theta_t + \Gamma X_t, \Sigma_\epsilon\right)$$

$$\Theta_t \sim \text{Gaussian}\left(\Phi\Theta_{t-1} + \Delta Z_t, \Sigma_\eta\right)$$

- Θ_t = states
- Θ_{t-1} = states (previous observation moment)
- Z_t = state covariates

Multivariate linear Gaussian state-space model

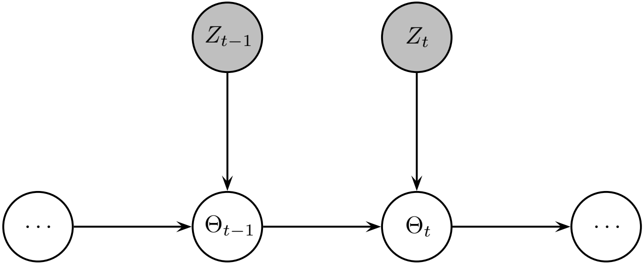
Transition equation

$$Y_t \sim \text{Gaussian}\left(\Psi\Theta_t + \Gamma X_t, \Sigma_\epsilon\right)$$

$$\Theta_t \sim \text{Gaussian}\left(\Phi\Theta_{t-1} + \Delta Z_t, \Sigma_\eta\right)$$

- Θ_t = states
- Θ_{t-1} = states (previous observation moment)
- Z_t = state covariates
- Φ = transition matrix (autocorrelations & cross-lagged relations)
- Δ = state covariate regression coefficient matrix
- Σ_η = innovation variance/covariance matrix

Multivariate linear Gaussian state-space model



Multivariate linear Gaussian state-space model

Observation equation

$$Y_t \sim \text{Gaussian}\left(\Psi\Theta_t + \Gamma X_t, \Sigma_\epsilon\right)$$

$$\Theta_t \sim \text{Gaussian}\left(\Phi\Theta_{t-1} + \Delta Z_t, \Sigma_\eta\right)$$

- Y_t = observations
- Θ_t = states
- X_t = observation covariates

Multivariate linear Gaussian state-space model

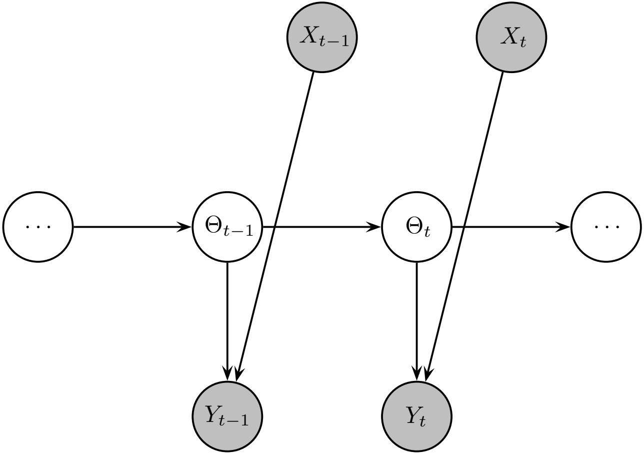
Observation equation

$$Y_t \sim \text{Gaussian}\left(\Psi\Theta_t + \Gamma X_t, \Sigma_\epsilon\right)$$

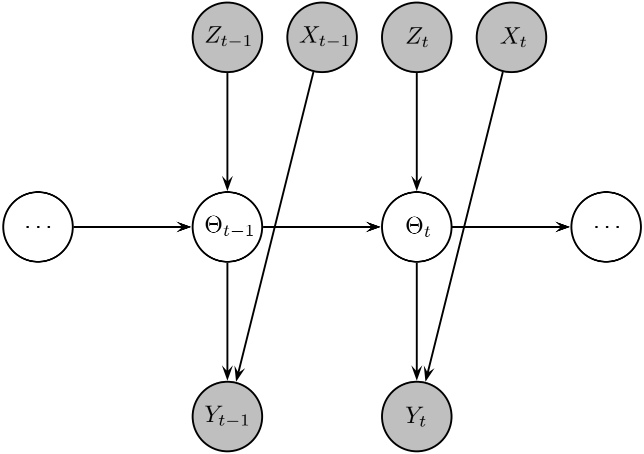
$$\Theta_t \sim \text{Gaussian}\left(\Phi\Theta_{t-1} + \Delta Z_t, \Sigma_\eta\right)$$

- Y_t = observations
- Θ_t = states
- X_t = observation covariates
- Ψ = design matrix (mapping $Y_t \rightarrow \Theta_t$)
- Γ = observation covariate regression coefficient matrix
- Σ_ϵ = observation error variance/covariance matrix

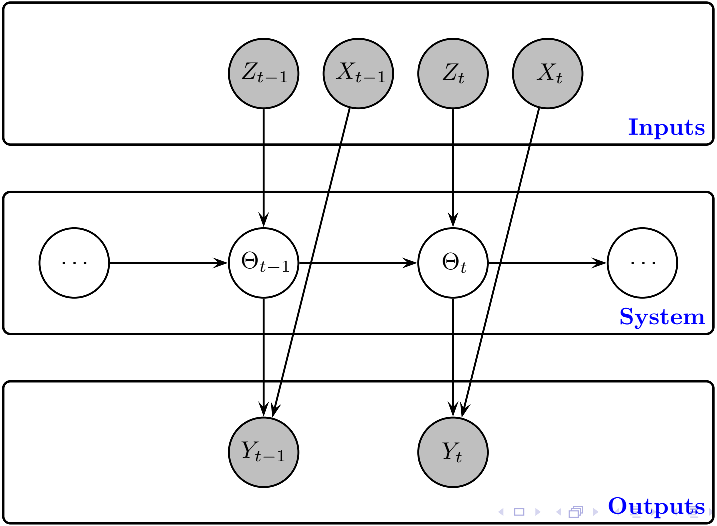
Multivariate linear Gaussian state-space model



Multivariate linear Gaussian state-space model



Multivariate linear Gaussian state-space model



Outline

Regime-switching MLGSS model

Regime-switching MLGSS model

$$Y_t \sim \text{Gaussian}\left(\Psi_t \Theta_t + \Gamma_t X_t, \Sigma_{\epsilon t}\right)$$

$$\Theta_t \sim \text{Gaussian}\left(\Phi_t \Theta_{t-1} + \Delta_t Z_t, \Sigma_{\eta t}\right)$$

Regime-switching MLGSS model

Regime-switching MLGSS model

$$Y_t \sim \text{Gaussian}\left(\Psi_t \Theta_t + \Gamma_t X_t, \Sigma_{\epsilon t}\right)$$
$$\Theta_t \sim \text{Gaussian}\left(\Phi_t \Theta_{t-1} + \Delta_t Z_t, \Sigma_{\eta t}\right)$$

$$\left(\begin{array}{c} \Psi_t, \Gamma_t, \Sigma_{\epsilon t} \\ \Phi_t, \Delta_t, \Sigma_{\eta t} \end{array}\right) = \begin{cases} \left(\begin{array}{c} \Psi_1, \Gamma_1, \Sigma_{\epsilon 1} \\ \Phi_1, \Delta_1, \Sigma_{\eta 1} \end{array}\right) & \text{if } R_t = 1 \\ \qquad \ldots & \\ \left(\begin{array}{c} \Psi_r, \Gamma_r, \Sigma_{\epsilon r} \\ \Phi_r, \Delta_r, \Sigma_{\eta r} \end{array}\right) & \text{if } R_t = r \end{cases}$$

- Dynamical properties $<$ current regime R_t
- Regime variable R_t observed or unobserved

Outline

Gibbs sampler

Gibbs sampler?

- Bayesian parameter estimation
- Iterative simulation algorithm
- Sampling distribution for each single parameter

Gibbs sampler

Gibbs sampler?

- Bayesian parameter estimation
- Iterative simulation algorithm
- Sampling distribution for each single parameter
- Convergence to joint posterior distribution
- Values simulate true joint posterior distribution
- Blocking efficient for some models

Gibbs sampler

Regime-switching MLGSS model

$$Y_t \sim \text{Gaussian}\left(\Psi_t \Theta_t + \Gamma_t X_t, \Sigma_\epsilon\right)$$

$$\Theta_t \sim \text{Gaussian}\left(\Phi_t \Theta_{t-1} + \Delta_t Z_t, \Sigma_\eta\right)$$

$$\begin{pmatrix} \Psi_t, \Gamma_t \\ \Phi_t, \Delta_t \end{pmatrix} = \begin{cases} \begin{pmatrix} \Psi_1, \Gamma_1 \\ \Phi_1, \Delta_1 \end{pmatrix} & \text{if } R_t = 1 \\ \ldots & \\ \begin{pmatrix} \Psi_r, \Gamma_r \\ \Phi_r, \Delta_r \end{pmatrix} & \text{if } R_t = r \end{cases}$$

Limitation: No regime-switching for Σ_ϵ & Σ_η

Gibbs sampler

Blocked Gibbs sampler

1. $\mathrm{p}(\Theta \mid \Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_\epsilon, \Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_\eta, Y, X, Z)$
2. $\mathrm{p}(\Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_\epsilon \mid \Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_\eta, \Theta, Y, X, Z)$
3. $\mathrm{p}(\Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_\eta \mid \Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_\epsilon, \Theta, Y, X, Z)$

Outline

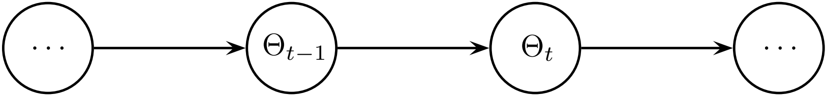
Sampling states

Blocked Gibbs sampler

1. $\mathrm{p}(\Theta \mid \Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_{\epsilon}, \Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_{\eta}, Y, X, Z)$
   $\longrightarrow$ Forward-filtering backward sampling
2. $\mathrm{p}(\Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_{\epsilon} \mid \Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_{\eta}, \Theta, Y, X, Z)$
3. $\mathrm{p}(\Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_{\eta} \mid \Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_{\epsilon}, \Theta, Y, X, Z)$

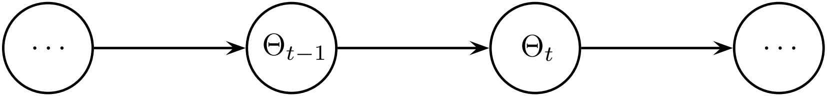
Sampling states

$\Rightarrow$ **Kalman filter** $\Rightarrow$



Sampling states

$\Rrightarrow$ **Kalman filter** $\Rrightarrow$



$\Lleftarrow$ **Backward-sampling** $\Lleftarrow$

Outline

Sampling parameters

Blocked Gibbs sampler

1. $p(\Theta \mid \Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_\epsilon, \Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_\eta, Y, X, Z)$
2. $p(\Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_\epsilon \mid \Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_\eta, \Theta, Y, X, Z)$
   $\longrightarrow$ Exact posterior simulation (decomposition)
3. $p(\Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_\eta \mid \Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_\epsilon, \Theta, Y, X, Z)$

Sampling parameters

Prior distributions

$$\Psi_{1,..,r} \propto 1$$

$$\Gamma_{1,..,r} \propto 1$$

$$\Sigma_\epsilon \sim \mathrm{W}^{-1}(\nu_0, S_0)$$

Sampling parameters

Prior distributions

$$\Psi_{1,..,r} \propto 1$$
$$\Gamma_{1,..,r} \propto 1$$
$$\Sigma_\epsilon \sim \mathrm{W}^{-1}(\nu_0, S_0)$$

Decomposition

$$\mathrm{p}(\Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_\epsilon \mid \Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_\eta, \Theta, Y, X, Z)$$

Sampling parameters

Prior distributions

$$\Psi_{1,..,r} \propto 1$$
$$\Gamma_{1,..,r} \propto 1$$
$$\Sigma_\epsilon \sim \mathrm{W}^{-1}(\nu_0, S_0)$$

Decomposition

$$\begin{aligned} & \mathrm{p}(\Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_\epsilon \mid \Phi_{1,..,r}, \Delta_{1,..,r}, \Sigma_\eta, \Theta, Y, X, Z) \\ = {} & \mathrm{p}(\Psi_{1,..,r}, \Gamma_{1,..,r}, \Sigma_\epsilon \mid \Theta, Y, X) \end{aligned}$$

Sampling parameters

Prior distributions

$$\Psi_{1,\ldots,r} \propto 1$$
$$\Gamma_{1,\ldots,r} \propto 1$$
$$\Sigma_\epsilon \sim \mathrm{W}^{-1}(\nu_0, S_0)$$

Decomposition

$$\begin{aligned}
&\mathrm{p}(\Psi_{1,\ldots,r}, \Gamma_{1,\ldots,r}, \Sigma_\epsilon \mid \Phi_{1,\ldots,r}, \Delta_{1,\ldots,r}, \Sigma_\eta, \Theta, Y, X, Z) \\
&= \mathrm{p}(\Psi_{1,\ldots,r}, \Gamma_{1,\ldots,r}, \Sigma_\epsilon \mid \Theta, Y, X) \\
&= \mathrm{p}(\Sigma_\epsilon \mid \Theta, Y, X) \rightarrow \text{sample from } \mathrm{W}^{-1} \\
&\times \mathrm{p}(\Psi_{1,\ldots,r} \mid \Theta, \Sigma_\epsilon, Y, X) \rightarrow \text{sample from Gaussian} \\
&\times \mathrm{p}(\Gamma_{1,\ldots,r} \mid \Theta, \Psi_{1,\ldots,r}, \Sigma_\epsilon Y, X) \rightarrow \text{sample from Gaussian}
\end{aligned}$$

Sampling parameters

Blocked Gibbs sampler

1. $p(\Theta \mid \Psi_{1,\ldots,r}, \Gamma_{1,\ldots,r}, \Sigma_\epsilon, \Phi_{1,\ldots,r}, \Delta_{1,\ldots,r}, \Sigma_\eta, Y, X, Z)$
2. $p(\Psi_{1,\ldots,r}, \Gamma_{1,\ldots,r}, \Sigma_\epsilon \mid \Phi_{1,\ldots,r}, \Delta_{1,\ldots,r}, \Sigma_\eta, \Theta, Y, X, Z)$
3. $p(\Phi_{1,\ldots,r}, \Delta_{1,\ldots,r}, \Sigma_\eta \mid \Psi_{1,\ldots,r}, \Gamma_{1,\ldots,r}, \Sigma_\epsilon, \Theta, Y, X, Z)$
   $\rightarrow$ Exact posterior simulation (decomposition)

Outline

Application

Data

- Oregon adolescent interaction study
- 13 depressed, 6 normal participants ($> 10\%$ anger)

Application

Data

- Oregon adolescent interaction study
- 13 depressed, 6 normal participants ($> 10\%$ anger)
- $Y_t = [\text{HR RSA SCL}]_t$

  $Z_t = [\text{FaAnger MoAnger}]_t$

  $R_t = [\text{AdAnger}]_t$

Application

Regime-switching MLGSS Model

- $Y_t \sim \text{Gaussian}\left(\Psi\Theta_t, \Sigma_\epsilon\right)$
  $\Theta_t \sim \text{Gaussian}\left(\Phi_t\Theta_{t-1} + \Delta Z_t, \Sigma_\eta\right)$
  $$\Phi_t = \begin{cases} \Phi_1 \text{ if } R_t = 1 \text{ (no anger)} \\ \Phi_2 \text{ if } R_t = 2 \text{ (anger)} \end{cases}$$
- *Fixed:* $\Psi = I_3$
- *Free:* $\Phi_1, \Phi_2, \Delta, \Sigma_\epsilon, \Sigma_\eta$

Application

$$\hat{\Delta}^{n} = \begin{array}{cc} & \begin{array}{cc} \text{FaAnger} & \text{MoAnger} \end{array} \\ \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} & \begin{bmatrix} 0 & 0 \\ -.01 & 0 \\ -.01 & 0 \end{bmatrix} \end{array} \qquad \hat{\Delta}^{d} = \begin{array}{cc} & \begin{array}{cc} \text{FaAnger} & \text{MoAnger} \end{array} \\ \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} & \begin{bmatrix} -.02 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \end{array}$$

$$\hat{\Phi}_{1}^{n} = \begin{array}{cc} \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} & \begin{bmatrix} .75 & .03 & .03 \\ -.03 & .75 & 0 \\ .01 & 0 & .98 \end{bmatrix} \end{array} \qquad \hat{\Phi}_{1}^{d} = \begin{array}{cc} \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} & \begin{bmatrix} .81 & 0 & .02 \\ -.05 & .71 & 0 \\ .01 & 0 & .98 \end{bmatrix} \end{array}$$

$$\hat{\Phi}_{2}^{n} = \begin{array}{cc} \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} & \begin{bmatrix} .74 & .03 & .02 \\ -.06 & .69 & -.01 \\ .04 & 0 & .98 \end{bmatrix} \end{array} \qquad \hat{\Phi}_{2}^{d} = \begin{array}{cc} \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} & \begin{bmatrix} .78 & .02 & .02 \\ -.02 & .67 & 0 \\ .02 & .03 & .98 \end{bmatrix} \end{array}$$

Application

$$\hat{\Delta}^n = \begin{array}{r} \\ \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{array}{c} \begin{array}{cc} \text{FaAnger} & \text{MoAnger} \end{array} \\ \begin{bmatrix} 0 & 0 \\ -.01 & 0 \\ -.01 & 0 \end{bmatrix} \end{array} \qquad \hat{\Delta}^d = \begin{array}{r} \\ \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{array}{c} \begin{array}{cc} \text{FaAnger} & \text{MoAnger} \end{array} \\ \begin{bmatrix} -.02 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \end{array}$$

$$\hat{\Phi}_1^n = \begin{array}{r} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .75 & .03 & .03 \\ -.03 & .75 & 0 \\ .01 & 0 & .98 \end{bmatrix} \qquad \hat{\Phi}_1^d = \begin{array}{r} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .81 & 0 & .02 \\ -.05 & .71 & 0 \\ .01 & 0 & .98 \end{bmatrix}$$

$$\hat{\Phi}_2^n = \begin{array}{r} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .74 & .03 & .02 \\ -.06 & .69 & -.01 \\ .04 & 0 & .98 \end{bmatrix} \qquad \hat{\Phi}_2^d = \begin{array}{r} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .78 & .02 & .02 \\ -.02 & .67 & 0 \\ .02 & .03 & .98 \end{bmatrix}$$

Application

$$\hat{\Delta}^{n} = \begin{array}{c|cc} & \text{FaAnger} & \text{MoAnger} \\ \hline \text{HR} & 0 & 0 \\ \text{RSA} & -.01 & 0 \\ \text{SCL} & -.01 & 0 \end{array} \qquad \hat{\Delta}^{d} = \begin{array}{c|cc} & \text{FaAnger} & \text{MoAnger} \\ \hline \text{HR} & -.02 & 0 \\ \text{RSA} & 0 & 0 \\ \text{SCL} & 0 & 0 \end{array}$$

$$\hat{\Phi}_1^{n} = \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .75 & .03 & .03 \\ -.03 & .75 & 0 \\ .01 & 0 & .98 \end{bmatrix} \qquad \hat{\Phi}_1^{d} = \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .81 & 0 & .02 \\ -.05 & .71 & 0 \\ .01 & 0 & .98 \end{bmatrix}$$

$$\hat{\Phi}_2^{n} = \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .74 & .03 & .02 \\ -.06 & .69 & -.01 \\ .04 & 0 & .98 \end{bmatrix} \qquad \hat{\Phi}_2^{d} = \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .78 & .02 & .02 \\ -.02 & .67 & 0 \\ .02 & .03 & .98 \end{bmatrix}$$

Application

$$\hat{\Delta}^{n} = \begin{array}{c|cc} & \text{FaAnger} & \text{MoAnger} \\ \hline \text{HR} & 0 & 0 \\ \text{RSA} & -.01 & 0 \\ \text{SCL} & -.01 & 0 \end{array} \qquad \hat{\Delta}^{d} = \begin{array}{c|cc} & \text{FaAnger} & \text{MoAnger} \\ \hline \text{HR} & -.02 & 0 \\ \text{RSA} & 0 & 0 \\ \text{SCL} & 0 & 0 \end{array}$$

$$\hat{\Phi}^{n}_{1} = \begin{array}{c|ccc} \text{HR} & .75 & .03 & .03 \\ \text{RSA} & -.03 & .75 & 0 \\ \text{SCL} & .01 & 0 & .98 \end{array} \qquad \hat{\Phi}^{d}_{1} = \begin{array}{c|ccc} \text{HR} & .81 & 0 & .02 \\ \text{RSA} & -.05 & .71 & 0 \\ \text{SCL} & .01 & 0 & .98 \end{array}$$

$$\hat{\Phi}^{n}_{2} = \begin{array}{c|ccc} \text{HR} & .74 & .03 & .02 \\ \text{RSA} & -.06 & .69 & -.01 \\ \text{SCL} & .04 & 0 & .98 \end{array} \qquad \hat{\Phi}^{d}_{2} = \begin{array}{c|ccc} \text{HR} & .78 & .02 & .02 \\ \text{RSA} & -.02 & .67 & 0 \\ \text{SCL} & .02 & .03 & .98 \end{array}$$

Application

$$\hat{\Delta}^{n} = \begin{array}{c} \\ \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{array}{cc} \text{FaAnger} & \text{MoAnger} \\ 0 & 0 \\ -.01 & 0 \\ -.01 & 0 \end{array} \qquad \hat{\Delta}^{d} = \begin{array}{c} \\ \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{array}{cc} \text{FaAnger} & \text{MoAnger} \\ -.02 & 0 \\ 0 & 0 \\ 0 & 0 \end{array}$$

$$\hat{\Phi}_1^{n} = \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .75 & .03 & .03 \\ -.03 & .75 & 0 \\ .01 & 0 & .98 \end{bmatrix} \qquad \hat{\Phi}_1^{d} = \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .81 & 0 & .02 \\ -.05 & .71 & 0 \\ .01 & 0 & .98 \end{bmatrix}$$

$$\hat{\Phi}_2^{n} = \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .74 & .03 & .02 \\ -.06 & .69 & -.01 \\ .04 & 0 & .98 \end{bmatrix} \qquad \hat{\Phi}_2^{d} = \begin{array}{c} \text{HR} \\ \text{RSA} \\ \text{SCL} \end{array} \begin{bmatrix} .78 & .02 & .02 \\ -.02 & .67 & 0 \\ .02 & .03 & .98 \end{bmatrix}$$

Outline

Conclusion

Summary

- Combination straightforward techniques
  $\longrightarrow$ Bayesian estimation algorithm for MLGSS model
- MLGSS model might be useful for investigation psycho(physio)logical processes
  - No differences found between groups / regimes (preliminary)
  - Model extensions might be necessary

Conclusion

Future work

- Proper prior distributions
- Regime-switching Σ_ϵ & Σ_η
- Hierarchical extension
- Non-Gaussian error distributions
- Model selection